(1) GENERAL INFORMATIC

(i) APPLICANT: ZEBEDEE, SUZANNE

INCHAUSPE, GENEVIEVE

NASOFF, MARC S.
PRINCE, ALFRED M.
HELTING, TORSTEN B.
DREVIN, HAKAN

DREVIN, HAKAN NUNN, MICHAEL F.

- (ii) TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING RECOMBINANT VIRAL ANTIGENS
 - (iii) NUMBER OF SEQUENCES: 29
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: James P. Hillman
 - (B) STREET: 45010 Pawnee Drive
 - (C) CITY: Fremont
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94539
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy Disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: Word Perfect 5.0 Dos Txt
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/931,855
 - (B) FILING DATE: Sep 16, 1997
 - (C) CLASSIFICATION: Cl. 424, subcl.192.1, 188.1 and 208.1
 - (C) CLASSIFICATION: Cl. 435, subcl.5 and Cl.422, subcl. 61
 - (C) CLASSIFICATION: Cl. 424, subcl.189.1, and 228.1
 - (C) CLASSIFICATION: Cl. 424, subcl.202.1
 - (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US08/563,733
 - (B) FILING DATE: 8-NOV-1995
 - (A) APPLICATION NUMBER: US08/049,531
 - (B) FILING DATE: 20-APR-1993
 - (A) APPLICATION NUMBER: US07/344,237
 - (B) FILING DATE: 26-APR-1989
 - (A) APPLICATION NUMBER: US07/191,229
 - (B) FILING DATE: 06-MAY-1988
 - (A) APPLICATION NUMBER: US07/206,499
 - (B) FILING DATE: 13-JUN-1988
 - (A) APPLICATION NUMBER: US07/258,016
 - (B) FILING DATE: 14-OCT-1988

- (A) APPLICAT N NUMBER: US08/272,271
- (B) FILING DATE: 8-JUL-1994
- (A) APPLICATION NUMBER: US07/616,369
- (B) FILING DATE: 21-NOV-1990
- (A) APPLICATION NUMBER: US07/573,643
- (B) FILING DATE: 27-AUG-1990

(viii) ATTORNEY/AGENT INFORMATION (0):

- (A) NAME: James P. Hillman Esq.
- (B) REGISTRATION NUMBER: 29748
- (C) REFERENCE/DOCKET NUMBER: 55467/69

(ix) TELECOMMUNICATION INFORMATION (O)

- (A) TELEPHONE: (510) 651 3991
- (B) TELEFAX: (510) 651 5991
- (C) TELEX:

(2)	IN (i)	SEQ (A (E	OUENC (1) (3)	LENG TYPE STRA	IARAC TH: : nu .NDEC	TERI 795 IClei NESS	c ac	S: pai id ngle									
	(ii)	MOI	ECUI	E TY	PE:	Geno	mic	DNA									
	(iii)	нуг	POTHE	ETICA	L: n	10											
	(iv)	PILA	ri-se	ENSE:	no												
	(ix)	(A	TURE () ()	NAME)S -789	ŀ									
	(xi)	SEÇ	QUENC	E DE	SCRI	PTIC	on: s	EQ I	D NC	:1:							
AGG	AGGGI	TT T	TCAT	TA T	G CC	A AT	C GT	G CA	G AA	C AI	C CA	.G GG	G C	AA A	rg g	TA	51
				Met	Pro) Ile	val	Gln 5	Asn	ıle	e Gln	Gly	Glr 10	n Met	Va]	L	
CAT	CAG	GCC	ATA	TCA	CCT	AGA	ACT	TTA	AAT	GC	TGG	GTA	AA	A GT	A GT	'A	99
His	Gln	Ala 15	Ile	Ser	Pro	Arg	Thr 20	Leu	Asn	Ala	Trp	Val 25	Lys	Val	Val		
GAA	GAG	AAG	GCT	TTC	AGC	CCA	GAA	GTG	ATA	CCC	ATG	TTI	TC	A GC	A TI	'A	147
Glu	Glu 30	Lys	Ala	Phe	Ser	Pro 35	Glu	Val	Ile	Pro	Met 40	Phe	Ser	Ala	Leu		
TCA	GAA	GGA	GCC	ACC	CCA	CAA	GAT	TTA	AAC	: ACC	C ATG	CTA	AA	C AC	A GI	'G	195
Ser 45	Glu	Gly	Ala	Thr	Pro 50	Gln	Asp	Leu	Asn	Thr 55	Met	Leu	Asn	Thr	Val 60		
GGG	GGA	CAT	CAA	GCA	GCC	ATG	CAA	ATG	TTA	AAA	A GAG	ACC	AT	C AA	r ga	.G	243
Gly	Gly	His	Gln	Ala 65	Ala	Met	Gln	Met	Leu 70	Lys	Glu	Thr	Ile	Asn 75	Glu		
GAA	GCT	GCA	GAA	TGG	GAT	AGA	GTG	CAT	CCA	GTO	CAT	GCA	GG	G CC	r at	T	291
Glu	Ala	Ala	Glu 80	Trp	Asp	Arg	Val	His 85	Pro	Val	His		Gly 90	Pro	Ile		
GCA	CCA	GGC	CAG	ATG	AGA	GAA	CCA	AGG	GGA	AGI	GAC	ATA	GC	A GG	A AC	T	339
Ala	Pro	Gly 95	Gln	Met	Arg	Glu	Pro 100	Arg	Gly	Ser	Asp	Ile 105	Ala	Gly	Thr		
ACT	AGT	ACC	CTT	CAG	GAA	CAA	ATA	GGA	TGG	ATO	ACA	LAA	' AA'	r cc	A CC	T	387
Thr	Ser 110	Thr	Leu	Gln	Glu	Gln 115	Ile	Gly	Trp	Met	Thr 120	Asn .	Asn	Pro	Pro		
ATC	CCA	GTA	GGA	GAA	ATT	TAT	AAA	AGA	TGG	ATA	ATC	CTG	GG.	A TT	A AA	T	435
Ile	Pro	Val	Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn		

AAA	ATA	GTA	AGA	ATG	TAT	AGC	CCT	ACC	AGC	ATT	CTG	GAC	ATA	AGA	CAA	483
Lys	Ile	Val	Arg	Met 145	Tyr	Ser	Pro		Ser 150	Ile :	Leu A	sp I		rg G 55	ln	
GGA	CCA	AAG	GAA	ccc	TTT	AGA	GAC	TAT	GTA	GAC	CGG	TTC	TAT	AAA	ACT	531
Gly	y Pro	Lys	5 Gli 160	ı Pro	Phe	a Arg	Asp	Tyr 165	Val	Asp	Arg		Tyr 70	Lys	Thr	
CTA	AGA	GCC	GAG	CAA	GCT	TCA	CAG	GAG	GTA	AAA	AAT	TGG	ATG	ACA	GAA	579
Leu	Arg	Ala 175	Glu	Gln	Ala	Ser	Gln 180	Glu	Val	Lys .	Asn I	rp M	et T	hr G	lu	
ACC	TTG	TTG	GTC	CAA	AAT	GCG	AAC	CCA	GAT	TGT	AAG	ACT	ATT	TTA	AAA	627
Thr	Leu 190	Leu	Val	Gln	Asn	Ala 195	Asn	Pro	Asp		Lys I 200	hr I	le L	eu L	ys	
GCA	TTG	GGA	CCA	GCG	GCT	ACA	CTA	GAA	GAA	ATG	ATG	ACA	GCA	TGT	CAG	675
Ala 205	Leu	Gly	Pro	Ala	Ala 210	Thr	Leu	Glu		Met : 215	Met 1	hr A	la C		ln 20	
GGA	GTA	GGA	GGA	ccc	AAA	AAT	CAA	CAA	TTA	TTA	TCC	TTA	TGG	GGG	TGT	723
Gly	Val	Gly	Gly	Pro 225	Lys	Asn	Gln		Leu 230	Leu	Ser I	Leu I		ly C 35	ys	
AAA	GGG	AAA	CTT	GTT	TGT	TAT	ACT	TCC	GTT	AAA	TGG	AAT	GGA	ccc	GGC	771
Lys	Gly	Lys	Leu 240	Val	Cys	Tyr	Thr	Ser 245	Val	Lys	Trp A		ly P 50	ro G	ly	
				GTT Val		TAA	TAA	79	5							
(2)	TN	TEAD!	<i>ለ</i> አመተ <i>ለ</i>	או דיר	אם פו	'O TE	NO.	2 •								

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 258 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile 5 10 15

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala 20 25 30

Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala 35 40 45

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln

. 60

Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu 105 Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly 125 115 120 Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg 135 Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu 150 155 145 Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu 170 Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val 180 185 190 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro 200 205 195 Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro Lys Asn Gln Gln Leu Leu Ser Leu Trp Gly Cys Lys Gly Lys Leu

Val Cys Tyr Thr Ser Val Lys Trp Asn Gly Pro Gly His Lys Ala Arg

Val Leu

225

- 50

(2) INFORMATION FOR SEQ ID NO:3:

245

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 795 base pairs
 - (B) TYPE: nucleic acid

230

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 16-789
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGGAGGGTTT TTCAT ATG CCA ATC GTG CAG AAC ATC CAG GGG CAA ATG GTA 51

Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val

5

235

250

240

255

CAT	CAG	GCC	ATA	TCA	CCT	AGA	ACT	TTA	LAA	GCA	A TGC	G GT	A AA	A G	TA	GTA	99
His	Gln	Ala 15	Ile	Ser	Pro	Arg	Thr 20	Leu	Asn	Ala	Trp	Val 25	Lys	Va]	L V	al	
GAA	GAG	AAG	GCT	TTC	AGC	CCA	GAA	GTG	ATA	CCC	C ATO	G TT	г тс	A G	CA	TTA	147
Glu	Glu 30	Lys	Ala	Phe	Ser	Pro 35	Glu	Val	Ile	Pro	Met 40	Phe	Ser	Ala	a L	eu	
TCA	GAA	GGA	GCC	ACC	CCA	CAA	GAT	TTA	AAC	ACC	ATC	G CT	A AA	C A	CA	GTG	195
Ser 45	Glu	Gly	Ala	Thr	Pro 50	Gln	Asp	Leu	Asn	Thr 55	Met	Leu	Asn	Thi	_	al O	
GGG	GGA	CAT	CAA	GCA	GCC	ATG	CAA	ATG	TTA	AAA	A GAG	G AC	C AT	C A	ΑТ	GAG	243
Gly	Gly	His	Gln	Ala 65	Ala	Met	Gln	Met	Leu 70	Lys	Glu	Thr	Ile	Asr 75	n G	lu	
GAA	GCT	GCA	GAA	TGG	GAT	AGA	GTG	CAT	CCA	GTC	CAT	r gc	A GG	G C	СТ	ATT	291
Glu	Ala	Ala	Glu 80	Trp	Asp	Arg	Val	His 85	Pro	Val	His	Ala	Gly 90	Pro) I	le	
GCA	CCA	GGC	CAG	ATG	AGA	GAA	CCA	AGG	GGA	AGI	GA(CAT	A GC	A G	GA	ACT	339
Ala	Pro	Gly 95	Gln	Met	Arg	Glu	Pro 100	Arg	Gly	Ser	Asp	Ile 105	Ala	Gly	<i>7</i> T	hr	
ACT	AGT	ACC	CTT	CAG	GAA	CAA	ATA	GGA	TGG	ATC	ACA	A AA	г аа	тс	CA	CCT	387
Thr	Ser 110	Thr	Leu	Gln	Glu	Gln 115	Ile	Gly	Trp	Met	Thr 120	Asn	Asn	Pro	P	ro	
ATC	CCA	GTA	GGA	GAA	ATT	TAT	AAA	AGA	TGG	ATA	A ATO	CT(G GG	АТ	TA	AAT	435
Ile 125	Pro	Val	Gly	Glu	Ile 130	Tyr	Lys	Arg	Trp	Ile 135	Ile	Leu	Gly	Let		sn 40	
AAA	ATA	GTA	AGA	ATG	TAT	AGC	CCT	ACC	AGC	: ATI	CTC	G GA	AT.	A A	GA	CAA	483
Lys	Ile	Val	Arg	Met 145	Tyr	Ser	Pro	Thr	Ser 150	Ile	Leu	Asp	Ile	Arc 155		ln	
GGA	CCA	AAG	GAA	ccc	TTT	AGA	GAC	TAT	GTA	GAC	CGC	G TT	TA	T A	AA	ACT	531
Gly	Pro	Lys	Glu 160	Pro	Phe	Arg	Asp	Туг 165	Val	Asp	Arg	Phe	Tyr 170	Lys	т	hr	
CTA	AGA	GCC	GAG	CAA	GCT	TCA	CAG	GAG	GTA	AAA	CAA A	TG(AT	G A	CA	GAA	579
Leu	Arg	Ala 175	Glu	Gln	Ala	Ser	Gln 180	Glu	Val	Lys	Asn	Trp 185	Met	Thr	- G	lu	
ACC	TTG	TTG	GTC	CAA	AAT	GCG	AAC	CCA	GAT	TGI	' AAC	ACT	T AT	T T	TA	AAA	627
Thr	Leu 190	Leu	Val	Gln	Asn	Ala 195	Asn	Pro	Asp	Cys	Lys 200	Thr	Ile	Leu	ı L	ys	
GCA	ттс	GGA	CCA	GCG	GCT	ACA	СТА	GAA	GAA	АТС	ATC	ACA	A GC	A T	GT	CAG	675

Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln 205 210 215 220

GGA GTA GGA GGA CCC AAA AAT CAA CAA AGA TTA AAT TTA TGG GGG TGT

Gly Val Gly Gly Pro Lys Asn Gln Gln Arg Leu Asn Leu Trp Gly Cys 225 230 235

AAA GGG AAA CTT ATT TGT TAT ACT TCC GTT AAA TGG AAT GGA CCC GGC 771

723

Lys Gly Lys Leu Ile Cys Tyr Thr Ser Val Lys Trp Asn Gly Pro Gly 240 245 250

CAT AAG GCA AGA GTT TTG TAA TAA 795 His Lys Ala Arg Val Leu

- (2) INFORMATION FOR SEQ ID NO:4:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 258 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile 5 10 15

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala 20 25 30

Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala 35 40 45

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln 50 55 60

Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu 65 70 75 80

Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln
85 90 95

Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu 100 105 110

Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly
115 120 125

Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg 130 135 140

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu 145 150 155 160

Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu 165 170 175

Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val 180 185 190

	Asn	Ala 195	Asn	Pro	Asp	ŗs	Lys 200	Thr	Ile	Leu	Lys	Ala 205	_eu	Gly	Pro	
Ala	Ala 210	Thr	Leu	Glu	Glu	Met 215	Met	Thr	Ala	Cys	Gln 220	Gly	Val	Gly	Gly	
Pro 225	Lys	Asn	Gln	Gln	Arg 230	Leu	Asn	Leu	Trp	Gly 235	Cys	Lys	Gly	Lys	Leu 240	
Ile	Cys	Tyr	Thr	Ser 245	Val	Lys	Trp	Asn	Gly 250	Pro	Gly	His	Lys	Ala 255	Arg	
Val	Leu															
(2)	II	1FORI	(ATI	ON FO	OR SE	Q II	NO:	5:								
	(i)	(<i>I</i> (I	-	LENC TYPE STRA	HARAC GTH: E: nu ANDEI DLOGY	795 Iclei NESS	base c ac	e pai cid ingle								
	(ii)	MOI	LECUI	LE TY	PE:	Geno	omic	DNA								
	(iii)	НУІ	POTHI	ETICA	AL: r	10										
	(iv)	INA	ri-si	ENSE	no											
	(ix)	(<i>I</i>	ATURI A) B)	NAM!	E/KEY ATION	T: 16	789									
AGG	(xi) AGGG				ESCRI G CC						C C	AG G	GG C	AA A	TG GT	A 51
				Met	Pro) Ile	e Val	Glr 5	n Ası	ı Ile	e Glr	ı Gly	7 Gl1 10	n Me	t Val	
CAT	CAG	GCC	ATA	TCA	CCT	AGA	ACT	TTA	CAA Z	GC2	A TGO	GT.	A AA	A GI	'A GTA	99
					CCT Pro											99
His	Gln	Ala 15	Ile	Ser	Pro	Arg	Thr 20	Leu	Asn	Ala	Trp	Val 25	Lys	Val		
His GAA	Gln GAG	Ala 15 AAG	Ile	Ser TTC	Pro	Arg CCA	Thr 20 GAA	Leu GTO	Asn ATA	Ala A CC	Trp	Val 25 G TT	Lys T TC	Val A GC	Val :A TTA	
His GAA Glu	Gln GAG Glu 30	Ala 15 AAG Lys	Ile GCT Ala	Ser TTC Phe	Pro AGC Ser	Arg CCA Pro 35	Thr 20 GAA Glu	Leu GTG Val	Asn ATA	Ala A CCC Pro	Trp C ATC Met 40	Val 25 G TT Phe	Lys T TC Ser	Val A GC Ala	Val :A TTA	
His GAA Glu TCA	Gln GAG Glu 30 GAA	Ala 15 AAG Lys GGA	Ile GCT Ala GCC	Ser TTC Phe	Pro AGC Ser	CCA Pro 35 CAA	Thr 20 GAA Glu	Leu GTG Val	Asn ; ATA Ile	Ala A CCO Pro	Trp Met 40	Val 25 G TT Phe	Lys T TC Ser A AA	Val A GC Ala C AC	Val A TTA Leu A GTG	147
His GAA Glu TCA Ser 45	Gln GAG Glu 30 GAA Glu	Ala 15 AAG Lys GGA Gly	Ile GCT Ala GCC Ala	Ser TTC Phe ACC Thr	Pro AGC Ser CCA Pro 50	CCA Pro 35 CAA Gln	Thr 20 GAA Glu GAI Asp	Leu GTG Val TTA	Asn Ile A AAC	Ala A CCC Pro C ACC Thr 55	Trp C ATC Met 40 C ATC	Val 25 G TT Phe G CT	Lys T TC Ser A AA Asn	Val A GC Ala C AC	Val Leu A GTG Val	147
His GAA Glu TCA Ser 45 GGG	Gln GAG Glu 30 GAA Glu GGA	Ala 15 AAG Lys GGA Gly CAT	Ile GCT Ala GCC Ala CAA	Ser TTC Phe ACC Thr	Pro AGC Ser CCA Pro 50	CCA Pro 35 CAA Gln	Thr 20 GAA Glu GAT Asp	Leu Val TTA Leu	Asn Ile AAAA Asn	Ala A CCC Pro C ACC Thr 55	Trp Met 40 CATC Met	Val 25 G TT Phe G CT Leu	Lys T TC Ser A AA Asn C AT	Val A GC Ala C AC Thr	Val Leu A GTG Val 60 T GAG	147 195
His GAA Glu TCA Ser 45 GGG Gly	Gln GAG Glu 30 GAA Glu GGA	Ala 15 AAG Lys GGA Gly CAT	GCT Ala GCC Ala CAA Gln	Ser TTC Phe ACC Thr GCA Ala 65	Pro AGC Ser CCA Pro 50 GCC Ala	CCA Pro 35 CAA Gln ATG	Thr 20 GAA Glu GAT Asp CAA	Leu Val TTA Leu ATG	Asn Ile AAAA Asn Leu 70	Ala A CCC Pro C ACC Thr 55 A AAA	Trp Met 40 A GAG	Val 25 G TT Phe Leu G AC	Lys T TC Ser A AA Asn C AT	Val A GC Ala C AC Thr C AA Asn 75	Val Leu A GTG Val 60 T GAG	147 195

GCA	CCA	GGC	CAG	ATG	AGA	AA	CCA	AGG	GGA	AGI	GAC	Al.	. GC	A GG	SA	ACT	339
Ala	Pro	Gly 95	Gln	Met	Arg	Glu	Pro 100	Arg	Gly	Ser	Asp	Ile 105	Ala	Gly	Tì	nr	
ACT	AGT	ACC	CTT	CAG	GAA	CAA	ATA	GGA	TGG	ATC	ACA	A AA	r aa	T CC	CA	CCT	387
Thr	Ser 110	Thr	Leu	Gln	Glu	Gln 115	Ile	Gly	Trp	Met	Thr 120	Asn	Asn	Pro	P	ro	
ATC	CCA	GTA	GGA	GAA	ATT	TAT	AAA	AGA	TGG	ATA	ATO	CTC	G GG	A TI	ſΆ	AAT	435
Ile 125	Pro	Val	Gly	Glu	Ile 130	Tyr	Lys	Arg	Trp	Ile 135	Ile	Leu	Gly	Leu		sn 40	
AAA	ATA	GTA	AGA	ATG	TAT	AGC	CCT	ACC	AGC	TTA :	CTC	G GA	C AT	A AC	ā	CAA	483
Lys	Ile	Val	Arg	Met 145	Tyr	Ser	Pro	Thr	Ser 150	Ile	Leu	Asp	Ile	Arg 155		ln	
GGA	CCA	AAG	GAA	ccc	TTT	AGA	GAC	TAT	GTA	GAC	CGG	TT(С ТА	T A	AA.	ACT	531
Gly	Pro	Lys	Glu 160	Pro	Phe	Arg	Asp	Tyr 165	Val	Asp	Arg	Phe	Tyr 170	Lys	T	hr	
CTA	AGA	GCC	GAG	CAA	GCT	TCA	CAG	GAG	GTA	AAA	AA	r TG	3 AT	G A	CA	GAA	579
Leu	Arg	Ala 175	Glu	Gln	Ala	Ser	Gln 180	Glu	Val	Lys	Asn	Trp 185	Met	Thr	G.	lu	
ACC	TTG	TTG	GTC	CAA	AAT	GCG	AAC	CCA	GAT	TGT	' AAC	AC!	r At	T TI	ľA	AAA	627
Thr	Leu 190	Leu	Val	Gln	Asn	Ala 195	Asn	Pro	Asp	Cys	Lys 200	Thr	Ile	Leu	L	ys	
GCA	TTG	GGA	CCA	GCG	GCT	ACA	CTA	GAA	GAA	ATC	AT(G AC	A GC	A TO	ЭT	CAG	675
	Leu	Gly	Pro	Ala	Ala 210	Thr	Leu	Glu	Glu	Met 215	Met	Thr	Ala	Cys	G:	ln 20	
205 GGA	GTA	GGA	GGA	CCA		AAT	CAA	CAA	CTI		AA	r TT	A TG	G GC			723
Gly	Val	Gly	Gly	Pro 225	Gln	Asn	Gln	Gln	Leu 230	Leu	Asn	Leu	Trp	Gly 235		ys	
AGA	GGG	AAA	GCT	ATT	TGT	TAT	ACT	TCC	GTI	CAA	YG(3 AA	r gg	A C	CC	GGC	771
Arg	Gly	Lys	Ala 240	Ile	Cys	Tyr	Thr	Ser 245	Val	Gln	Trp	Asn	Gly 250	Pro	G:	ly	
_				GTT Val		TAA	TAA	79	5								

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 258 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Pro Ile Val Gln Asn ile Gln Gly Gln Met Val His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala 25 Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln 55 Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu 65 Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu 105 Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly 125 120 115 Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg 135 Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu 150 155 Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu 165 Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val 180 185 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro 200 205 Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly 210 Pro Gln Asn Gln Gln Leu Leu Asn Leu Trp Gly Cys Arg Gly Lys Ala 230 235 225 Ile Cys Tyr Thr Ser Val Gln Trp Asn Gly Pro Gly His Lys Ala Arg 250

Val Leu

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 378 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (iii) HYPOTHETICAL: no

ļ	(ix)	FEA (A	•	NAME	E/KEY ATION)S 5-375	i									
	(xi)	SEC	QUENC	E DE	ESCRI	PTIC	N: S	EQ I	D NC	:7:							
AGG?	AGGGI	TT I	rtca:	TA T	G AG	C AC	G AA	т сс	T AA	A CC	T CA	A A	SA A	AA A	CC A	AA	51
				Met	: Ser	Thr	. Asr	Pro 5	Lys	Pro	Gln	Arg	Lys 10	s Thi	c Lys	S	
CGT	AAC	ACC	AAC	CGT	CGC	CCA	CAG	GAC	GTC	AAG	TTC	ccc	G GG	T GG	C GG	T	99
Arg	Asn	Thr 15	Asn	Arg	Arg	Pro	Gln 20	Asp	Val	Lys	Phe	Pro 25	Gly	Gly	Gly		
CAG	ATC	GTT	GGT	GGA	GTT	TAC	TTG	TTG	cce	CGC	AGC	GG	e cc	T AG	TT A	ľG	147
Gln	Ile 30	Val	Gly	Gly	Val	Tyr 35	Leu	Leu	Pro	Arg	Arg 40	Gly	Pro	Arg	Leu		
GGT	GTG	CGC	GCG	ACG	AGG	AAG	ACT	TCC	GAG	CGG	TCG	CA	A CC	T CG	A GG	ST	195
Gly 45	Val	Arg	Ala	Thr	Arg 50	Lys	Thr	Ser	Glu	Arg 55	Ser	Gln	Pro	Arg	Gly 60		
AGA	CGT	CAG	CCT	ATC	ccc	AAG	GTG	CGT	CGG	CCG	GAG	G GG	C AG	G AC	C TG	€G	243
Arg	Arg	Gln	Pro	Ile 65	Pro	Lys	Val	Arg	Arg 70	Pro	Glu	Gly	Arg	Thr 75	Trp		
GCT	CAG	CCC	GGG	TAC	CCT	TGG	ccc	CTC	LAT	' GGC	CAA C	GA(G GG	T TG	C GG	€G	291
Ala	Gln	Pro	Gly 80	Tyr	Pro	Trp	Pro	Leu 85	Tyr	Gly	Asn	Glu	Gly 90	Cys	Gly		
TGG	GCG	GGA	TGG	CTC	CTG	TCT	ccc	CGT	GGC	TCI	CGG	CC!	r ag	C TG	G GG	€C	339
Trp	Ala	Gly 95	Trp	Leu	Leu	Ser	Pro 100	Arg	Gly	Ser	Arg	Pro 105	Ser	Trp	Gly		
			CCC Pro									TAA	31	78			
(2)	I	NFORI	MATIC	ON FO	OR SE	Q II	NO:	8:									
	(i)	(<i>I</i> (I	в)	TYPE TOPO	TH: E: am OLOGY	120 nino : li	amir acid near	no ac l	ids								
	(ii) (xi) Ser	SEÇ	QUENC Asn	E DE	ESCRI	PTIC	N: S				Lys	Arg	Asn	Thr 15	Asn		
Arg	Arg	Pro	Gln 20	Asp	Val	Lys	Phe	Pro 25	Gly	Gly	Gly	Gln	Ile 30	Val	Gly		

(iv) ANTI-SENSE: no

	GIA	val	35	Leu	Leu	Pro	.g	40	GIY	PIO	Arg	Leu	45	aı	AIG	Ald	
•	Thr	Arg 50	Lys	Thr	Ser	Glu	Arg 55	Ser	Gln	Pro	Arg	Gly 60	Arg	Arg	Gln	Pro	
	Ile 65	Pro	Lys	Val	Arg	Arg 70	Pro	Glu	Gly	Arg	Thr 75	Trp	Ala	Gln	Pro	Gly 80	
	Tyr	Pro	Trp	Pro	Leu 85	Tyr	Gly	Asn	Glu	Gly 90	Cys	Gly	Trp	Ala	Gly 95	Trp	
	Leu	Leu	Ser	Pro 100	Arg	Gly	Ser	Arg	Pro 105	Ser	Trp	Gly	Pro	Thr 110	Asp	Pro	
	Arg (2)	_	115				Leu EQ II	120	9:			٠					
	1	(i)	(2 (1 (0	QUENCA) 3) C) O)	LENC TYPI STRA	STH: E: nu ANDEI	TERI 378 aclei ONESS	base ic ac : si	e pai cid ingle								
		(ii)	MOI	LECUI	LE TY	PE:	Gend	omic	DNA								
		(iii)	НУІ	POTHI	ETICA	AL: 1	no										
		(iv)	ANT	ri-si	ENSE	no											
	((ix)	(2	ATURI A) 3)	NAMI		7: CI N: 16		5								
		(xi) AGGGT					EPTIC C AC					CT C	AA A	GA A	AA A	.CC AAA	51
					Met	Sei	c Thi	: Asr	Pro 5	Lys	s Pro	o Glr	n Arg	J Ly:	s Thi	r Lys	
	CGT	AAC	ACC	AAC	CGT	CGC	CCA	CAG	GAC	GTO	CAAG	G TT	C CC	G GG	T GG	C GGT	99
	Arg	Asn	Thr 15	Asn	Arg	Arg	Pro	Gln 20	Asp	Val	Lys	Phe	Pro 25	Gly	Gly	Gly	
	CAG	ATC	GTT	GGT	GGA	GTT	TAC	TTG	TTC	cco	G CG	C AG	G GG	c cc	T AG	A TTG	147
	Gln	Ile 30	Val	Gly	Gly	Val	Tyr 35	Leu	Leu	Pro	Arg	Arg 40	Gly	Pro	Arg	Leu	
	GGT	GTG	CGC	GCG	ACG	AGG	AAG	ACT	TCC	GAC	G CG	G TC	G CA	A CC	T CG	A GGT	195
	Gly 45	Val	Arg	Ala	Thr	Arg 50	Lys	Thr	Ser	Glu	Arg 55	Ser	Gln	Pro	Arg	Gly 60	
	AGA	CGT	CAG	CCT	ATC	ccc	AAG	GCA	CGI	CGG	G CC	C GA	G GG	C AG	G AC	C TGG	243
	Arg	Arg	Gln	Pro	Ile 65	Pro	Lys	Ala	Arg	Arg 70	Pro	Glu	Gly	Arg	Thr 75	Trp	
	GCT	CAG	ccc	GGG	TAC	CCT	TGG	ccc	CTC	TAT	r GG	C AA	r ga	G GG	T TG	C GGG	291

Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly

TGG GCG GGA TGG CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC

Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly 95

CCC ACA GAC CCC CGG CGT AGG TCG CGC AAT TTG GGT TAA 378 Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly 115 110

- INFORMATION FOR SEQ ID NO:10:
 - SEQUENCE CHARACTERISTICS:
 - LENGTH: 120 amino acids
 - TYPE: amino acid (B)
 - TOPOLOGY: linear (D)
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala 35

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro

Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly 80 70 75

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp 85

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro 110 105 100

Arg Arg Arg Ser Arg Asn Leu Gly 115

- INFORMATION FOR SEQ ID NO:11: (2)
 - SEQUENCE CHARACTERISTICS: (i)
 - LENGTH: 378 base pairs
 - (B) TYPE: nucleic acid
 - STRANDEDNESS: single (C)
 - TOPOLOGY: linear (D)
 - MOLECULE TYPE: Genomic DNA (ii)
 - (iii) HYPOTHETICAL: no
 - ANTI-SENSE: no (iv)
 - FEATURE: (ix)

339

(A) (B)	NAME/KE' CDS LOCATION: 16-375	
(xi) SEQUEN	ICE DESCRIPTION: SEQ ID NO:11:	
AGGAGGGTTT TTCA	AT ATG AGC ACG AAT CCT AAA CCT CAA AGA AAA ACC AAA	51
	Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys 5 10	
CGT AAC ACC AAC	C CGT CGC CCA CAG GAC GTC AAG TTC CCG GGT GGC GGT	99
Arg Asn Thr Asn 15	Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly 20 25	
CAG ATC GTT GGT	I GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG 14	47
Gln Ile Val Gly 30	Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu 35 40	
GGT GTG CGC GCG	G ACG AGG AAG ACT TCC GAG CGG TCG CAA CCT CGA GGT 19	95
Gly Val Arg Ala 45	Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly 50 55 60	
AGA CGT CAG CCT	T ATC CCC AAG GAC CGT CGG TCC ACG GGC AAG TCC TGG 24	43
Arg Arg Gln Pro	o Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp 65 70 75	
GGT AAG CCC GGG	G TAC CCT TGG CCC CTC TAT GGC AAT GAG GGT TGC GGG 29	91
Gly Lys Pro Gly 80	Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly 85 90	
TGG GCG GGA TGG	G CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC 33	39
Trp Ala Gly Trp 95	Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly 100 105	
	C CGG CGT AGG TCG CGC AAT TTG GGT TAA 378 O Arg Arg Arg Ser Arg Asn Leu Gly 115 120	
(2) INFORMATI	ON FOR SEQ ID NO:12:	
(i) SEQUEN (A) (B) (D)	ICE CHARACTERISTICS: LENGTH: 120 amino acids TYPE: amino acid TOPOLOGY: linear	
(ii) MOLECU	JLE TYPE: protein	
(xi) SEQUEN	ICE DESCRIPTION: SEQ ID NO:12:	
Met Ser Thr Asn	Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn 5 10 15	
Arg Arg Pro Gln 20	Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly 25 30	

Gly Val Tyr Leu Leu Pro Arg Gly Pro Arg Leu Gly Val Arg Ala

40 45

35

Thr	Arg 50	Lys	Thr	Ser	Glu	Arg 55	Ser	Gln	Pro	Arg	Gly 60	Arg	Arg	Gln	Pro	
Ile 65	Pro	Lys	Asp	Arg	Arg 70	Ser	Thr	Gly	Lys	Ser 75	Trp	Gly	Lys	Pro	Gly 80	
	Pro	Trp	Pro	Leu 85		Gly	Asn	Glu	Gly 90		Gly	Trp	Ala	Gly 95	Trp	
Leu	Leu	Ser	Pro 100	Arg	Gly	Ser	Arg	Pro 105	Ser	Trp	Gly	Pro	Thr 110	Asp	Pro	
Arg	Arg	Arg 115	Ser	Arg	Asn	Leu	Gly 120									
(2)	II	VFORI	OITAN	ON FO	OR SI	EQ II	ом с	13:								
	(i)	(2 (1 (0	QUENCA) 3) C) O)	LENG TYPI STRA	HARAG ETH: E: nu ANDEI DLOGY	378 iclei ONESS	base ic ac	e pai cid ingle								
	(ii)	MOI	LECUI	LE TY	PE:	Gend	omic	DNA								
	(iii)	HYI	POTHI	ETIC	AL: r	10										
	(iv)	ANT	ri-si	ENSE	no											
	(ix)	(Z	ATURI A) 3)	NAM	E/KEY ATION			5								
	(xi)	SEÇ	QUENC	CE DI	ESCRI	PTIC	ON: S	SEQ 1	D NO	13:	:					
AGG	AGGGT	TT T	TTCA:	r at	G AG	C AC	G AA	T CC	T AA	A CO	CT C	AA A	GA A	AA A	CC AAA	51
				Met	Ser	Thr	: Asr	Pro 5	Lys	s Pro	o Glr	n Arg	10	s Thi	r Lys	
CGT	AAC	ACC	AAC	CGT	CGC	CCA	CAG	GAC	GTO	C AAC	G TT	c cc	G GG	T GG	C GGT	99
Arg	Asn	Thr 15	Asn	Arg	Arg	Pro	Gln 20	Asp	Val	Lys	Phe	Pro 25	Gly	Gly	Gly	
CAG	ATC	GTT	GGT	GGA	GTT	TAC	TTG	TTG	CCG	G CGC	C AG	G GG	c cc	T AG	A TTG	147
Gln	Ile 30	Val	Gly	Gly	Val	Tyr 35	Leu	Leu	Pro	Arg	Arg 40	Gly	Pro	Arg	Leu	
GGT	GTG	CGC	GCG	ACG	AGG	AAG	ACT	TCC	GAG	CGC	TC	G CA	A CC	T CG	A GGT	195
Gly 45	Val	Arg	Ala	Thr	Arg 50	Lys	Thr	Ser	Glu	Arg 55	Ser	Gln	Pro	Arg	Gly 60	
AGA	CGT	CAG	CCT	ATC	CCC	AAG	GCA	CGT	CGG	TCC	C GAG	G GG	C AG	G TC	C TGG	243
Arg	Arg	Gln	Pro		Pro	Lys	Ala	Arg		Ser	Glu	Gly	Arg		Trp	
GCT	CAG	ccc	GGG	65 TAC	CCT	TGG	CCC	CTC	70 TAI	' GGC	CAA	GA(G GG'	75 T TG	C GGG	291

Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly 80 85 90

TGG GCG GGA TGG CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC 339

Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly 95 100 105

CCC ACA GAC CCC CGG CGT AGG TCG CGC AAT TTG GGT TAA 378
Pro Thr Asp Pro Arg Arg Ser Arg Asn Leu Gly
110 115 120

- (2) INFORMATION FOR SEQ ID NO:14:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 120 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gln Ile Val Gly 20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala 35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro 50 55 60

Ile Pro Lys Ala Arg Arg Ser Glu Gly Arg Ser Trp Ala Gln Pro Gly 65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp
85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro 100 105 110

Arg Arg Arg Ser Arg Asn Leu Gly
115 120

- (2) INFORMATION FOR SEQ ID NO:15:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 381 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: Genomic DNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS

(B) LOCATIO: 16-375

	(xi) AGGGT		QUENC PTCAT									AT C	AT G	GC C	CG	GGC	51
				Met	Pro	Ile	e His	s His 5	s His	His	His	s His	Gly 10	y Pro	o G	ly	
TCC	GTC	ACT	GTG	TCC	CAT	CCT	AAC	: ATC	GAG	GAG	GT	r GC	r ct	G TC	CC A	/CC	99
Ser	Val	Thr 15	Val	Ser	His	Pro	Asn 20	Ile	Glu	Glu	Val	Ala 25	Leu	Ser	Th:	r	
ACC	GGA	GAG	ATC	ccc	TTT	TAC	GGC	: AAG	GCI	TATO	c cc	C CT	C GA	G GI	rg A	ATC	147
Thr	Gly 30	Glu	Ile	Pro	Phe	Tyr 35	Gly	Lys	Ala	Ile	Pro 40	Leu	Glu	Val	Ile	е	
AAG	GGG	GGA	AGA	CAT	CTC		TTC	TGC	CAC	TC		G AA	G AA	G TG	C G	SAC	195
Lys 45	Gly	Gly	Arg	His	Leu 50	Ile	Phe	Cys	His	Ser 55	Lys	Lys	Lys	Cys	As) 60		
GAG	CTC	GCC	GCG	AAG	CTG	GTC	GCA	TTC	GGG	ATC	AA!	r gc	C GT	G GC	CC I	TAC	243
Glu	Leu	Ala	Ala	Lys 65	Leu	Val	Ala	Leu	Gly 70	Ile	Asn	Ala	Val	Ala 75	Ту	r	
TAC	CGC	GGT	CTT	GAC	GTG	TCT	GTC	ATC	ccc	ACC	C AG	C GG	C GA	T GI	T C	STC	291
Tyr	Arg	Gly	Leu 80	Asp	Val	Ser	Val	Ile 85	Pro	Thr	Ser	Gly	Asp 90	Val	Va:	1	
GTC	GTG	TCA	ACC	GAT	GCT	CTC	ATG	ACI	GGC	TTI	r Ac	C GG	C GA	C TI	C G	SAC	339
Val	Val	Ser 95	Thr	Asp	Ala	Leu	Met 100	Thr	Gly	Phe	Thr	Gly 105	Asp	Phe	Asj	p	
	GTG Val 110												TAA	38	81		
(2)	IN	FORM	OITAN	N FO	R SE	Q II	NO:	16:									
ı	(i)	(Z	•	LENG		121 ino	amir acid	no ac I	cids								
((ii)	MOI	LECUI	E TY	PE:	prot	ein										
	(xi)	SEÇ	QUENC	E DE	SCRI	PTIC	N: S	SEQ I	D NC	:16:							
Met	Pro	Ile	His	His 5	His	His	His	His	Gly 10	Pro	Gly	Ser	Val	Thr 15	Va:	1	
Ser	His	Pro	Asn 20	Ile	Glu	Glu	Val	Ala 25	Leu	Ser	Thr	Thr	Gly 30	Glu	Ile	9	
Pro	Phe	Tyr 35	Gly	Lys	Ala	Ile	Pro 40	Leu	Glu	Val	Ile	Lys 45	Gly	Gly	Ar	3	

His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala

60

	•														
·Lys	Leu	Val	Ala	Leu	Gly	Ile	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly	Leu

Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val Val Ser Thr 85 90 95

Asp Ala Leu Met Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp 100 105 110

Cys Asn Thr Gly Thr Glu Leu Glu Phe 115 120

- (2) INFORMATION FOR SEQ ID NO:17:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 774 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: Genomic DNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 16-771
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
 AGGAGGGTTT TTCAT ATG TCC CCT ATA CTA GGT TAT TGG AAA ATT AAG GGC 51

Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly
5

- CTT GTG CAA CCC ACT CGA CTT CTT TTG GAA TAT CTT GAA GAA AAA TAT 99
- Leu Val Gln Pro Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr
 15 20 25
- GAA GAG CAT TTG TAT GAG CGC GAT GAA GGT GAT AAA TGG CGA AAC AAA 147
- Glu Glu His Leu Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys 30 35 40
- AAG TTT GAA TTG GGT TTG GAG TTT CCC AAT CTT CCT TAT TAT ATT GAT 195
- Lys Phe Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp 50 55 60
- GGT GAT GTT AAA TTA ACA CAG TCT ATG GCC ATC ATA CGT TAT ATA GCT 243
- Gly Asp Val Lys Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala
 65 70 75
- GAC AAG CAC AAC ATG TTG GGT GGT TGT CCA AAA GAG CGT GCA GAG ATT 291
- Asp Lys His Asn Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile 80 85 90

TCA.	ATG	CTT	GAA	GGA	GCC	'TT	TTG	GAT	LLV	AGA	TAC	C GC	GT'	T TC	G AG	A 339
Ser	Met	Leu 95	Glu	Gly	Ala	Val	Leu 100	Asp	Ile	Arg	Tyr	Gly 105	Val	Ser	Arg	
ATT	GCA	TAT	AGT	AAA	GAC	TTT	GAA	ACT	CTC	AAA	GT	C GA	r TT	T CT	r AG	C 387
Ile	Ala 110	Tyr	Ser	Lys	Asp	Phe 115	Glu	Thr	Leu	Lys	Val 120	Asp	Phe	Leu	Ser	
AAG	CTA	CCT	GAA	ATG	CTG	AAA	ATG	TTC	GAA	A GAT	CG	r T T	A TG	T CA	r aa	A 435
Lys 125	Leu	Pro	Glu	Met	Leu 130	Lys	Met	Phe	Glu	Asp 135	Arg	Leu	Cys	His	Lys 140	
ACA	TAT	TTA	AAT	GGT	GAT	CAT	GTA	ACC	CAT	CCI	GAC	C TT	C AT	G TT	G TA	T 483
Thr	Tyr	Leu	Asn	Gly 145	Asp	His	Val	Thr	His 150	Pro	Asp	Phe	Met	Leu 155	Tyr	
GAC	GCT	CTT	GAT	GTT	GTT	TTA	TAC	ATG	GAC	CCA	ATO	G TG	C CT	G GA	r gc	G 531
Asp	Ala	Leu	Asp 160	Val	Val	Leu	Tyr	Met 165	Asp	Pro	Met	Cys	Leu 170	Asp	Ala	
TTC	CCA	AAA	TTA	GTT	TGT	TTT	AAA	AAA	CGI	TTA T	GA/	A GC	T AT	c cc	A CA	A 579
Phe	Pro	Lys 175	Leu	Val	Cys	Phe	Lys 180	Lys	Arg	Ile	Glu	Ala 185	Ile	Pro	Gln	
ATT	GAT	AAG	TAC	TTG	AAA	TCC	AGC	AAG	TAT	TATA	GC?	A TG	G CC	T TT	G CA	G 627
Ile	Asp 190	Lys	Tyr	Leu	Lys	Ser 195	Ser	Lys	Tyr	Ile	Ala 200	Trp	Pro	Leu	Gln	
GGC		CAA	GCC	ACG	TTT		GGT	GGC	GAC	CAT		r cc	A AA	A TC	G GA	T 675
Gly 205	_	Gln	Ala	Thr		Gly			Asp	His 215	Pro	Pro	Lys	Ser	Asp 220	
CTG	GTT	CCG	CGT	GGA	TCC	GAC	GTC	AAG	TTC	CCG	GGT	r GG	C GG	T CA	G AT	C 723
Leu	Val	Pro	Arg	Gly 225	Ser	Asp	Val	Lys	Phe 230	Pro	Gly	Gly	Gly	Gln 235	Ile	
GTT	GGT	GGA	GTT	TAC	TTG	TTG	CCG	CGC	AGG	GAA	A TTC	C AT	C GT	G AC	r ga	C 771
Val	Gly	Gly	Val 240	Tyr	Leu	Leu	Pro	Arg 245	Arg	Glu	Phe	Ile	Val 250	Thr	Asp	

TGA 774

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 252 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18: Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro

5

Thr Arg Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu 25 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys 50 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser 105 100 Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu 120 115 Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp 145 150 155 160 Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr 185 Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala 200 205 Thr Phe Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg 220 Gly Ser Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val 225 240

Tyr Leu Leu Pro Arg Arg Glu Phe Ile Val Thr Asp 245 250

INFORMATION FOR SEQ ID NO:19: (2)

- (i) SEQUENCE CHARACTERISTICS:
 - LENGTH: 31 base pairs (A)
 - TYPE: nucleic acid (B)
 - STRANDEDNESS: single (C)
 - TOPOLOGY: linear (D)
- MOLECULE TYPE: Genomic DNA (ii)
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- SEQUENCE DESCRIPTION: SEQ ID NO:19: (xi)

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GACCCGGCCA TAAGGCAAGA GTTTTGTAAT AAG 33

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: yes
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21: GATCCTTATT ACAAAACTCT TGCCTTATGG CCGG 34
- (2) INFORMATION FOR SEQ ID NO:22:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: Genomic DNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCTCGCATAT GAGCACGATT CCCAAACC 28

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEL ESS: single TOPOLOGY: linear (D) MOLECULE TYPE: Genomic DNA (ii) (iii) HYPOTHETICAL: no ANTI-SENSE: yes (iv) SEQUENCE DESCRIPTION: SEQ ID NO:23: (xi) 32 GACGAATTCT TAACCCAAAT TGCGCGACCT AC INFORMATION FOR SEQ ID NO:24: (2) SEQUENCE CHARACTERISTICS: (i) LENGTH: 66 base pairs (A) TYPE: nucleic acid (B) (C) STRANDEDNESS: single TOPOLOGY: linear (D) (ii) MOLECULE TYPE: Genomic DNA (iii) HYPOTHETICAL: no (iv) ANTI-SENSE: no SEQUENCE DESCRIPTION: SEQ ID NO:24: (xi) GATCCGACGT CAAGTTCCCG GGTGGCGGTC AGATCGTTGG TGGAGTTTAC TTGTTGCCGC 60 GCAGGG 66 INFORMATION FOR SEQ ID NO:25: (2) SEQUENCE CHARACTERISTICS: (i) LENGTH: 66 base pairs (A) TYPE: nucleic acid (B) STRANDEDNESS: single (C) TOPOLOGY: linear (D) (ii) MOLECULE TYPE: Genomic DNA (iii) HYPOTHETICAL: no ANTI-SENSE: yes (iv) SEQUENCE DESCRIPTION: SEQ ID NO:25: (xi) AATTCCCTGC GCGGCAACAA GTAAACTCCA CCAACGATCT GACCGCCACC CGGGAACTTG 60 ACGTCG 66 INFORMATION FOR SEQ ID NO:26: (2)

(i)

SEQUENCE CHARACTERISTICS:

· .	(A) LENGTH: 3 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
(ii)	MOLECULE TYPE: Genomic DNA
(iii)	HYPOTHETICAL: no
(iv)	ANTI-SENSE: no
(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:26:
GGAATTCCAT ATGTCCCCTA TACTAGGT 28	
(2) IN	FORMATION FOR SEQ ID NO:27:
(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 26 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
(ii)	MOLECULE TYPE: Genomic DNA
(iii)	HYPOTHETICAL: no
(iv)	ANTI-SENSE: yes
(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:27:
CGGAATTCTC ACCTGCGCGG CAACAA 26	
(2) IN	FORMATION FOR SEQ ID NO:28:
(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 52 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
(ii)	MOLECULE TYPE: Genomic DNA
(iii)	HYPOTHETICAL: no
(iv)	ANTI-SENSE: no
(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:28:
TATGCCTA	TT CATCATCATC ATCATCATGG CCCGGGAATT CTAAGTAAGT AG 52
(2) INFORMATION FOR SEQ ID NO:29:	
(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 54 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
(ii)	MOLECULE TYPE: Genomic DNA
(iii)	HYPOTHETICAL: no

- (iv) ANTI-SENSE: ye.
- (xí) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCCTACTT ACTTAGAATT CCCGGGCCAT GATGATGATG ATGATGAATA GGCA 54 ^Z